**МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И МАССОВЫХ**

**КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение

высшего образования

**«Поволжский государственный университет телекоммуникаций и информатики»**

Факультет ***Информационных систем и технологий***

Направление подготовки ***09.04.02 Информационные системы и технологии***

**Отчет по лабораторной работе №7**

**По дисциплине “Системы и методы искусственного интеллекта”**

**студента 1 курса группы**

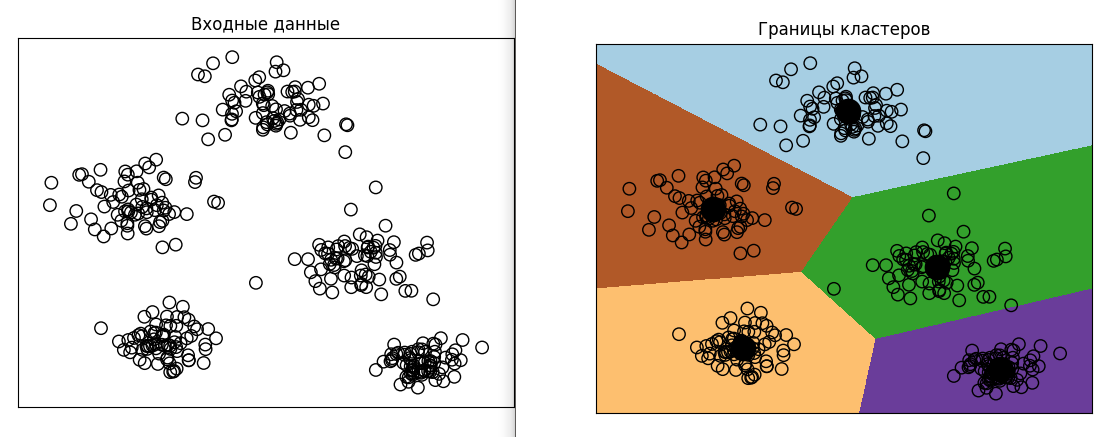
**\_\_**ИСТм-43**\_\_\_**

Емелин Максим Сергеевич

(фамилия, имя, отчество)

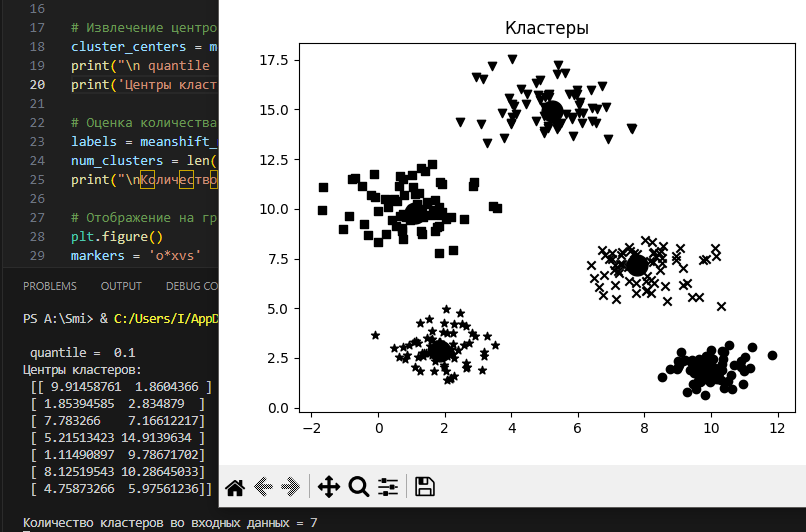
Самара, 2024

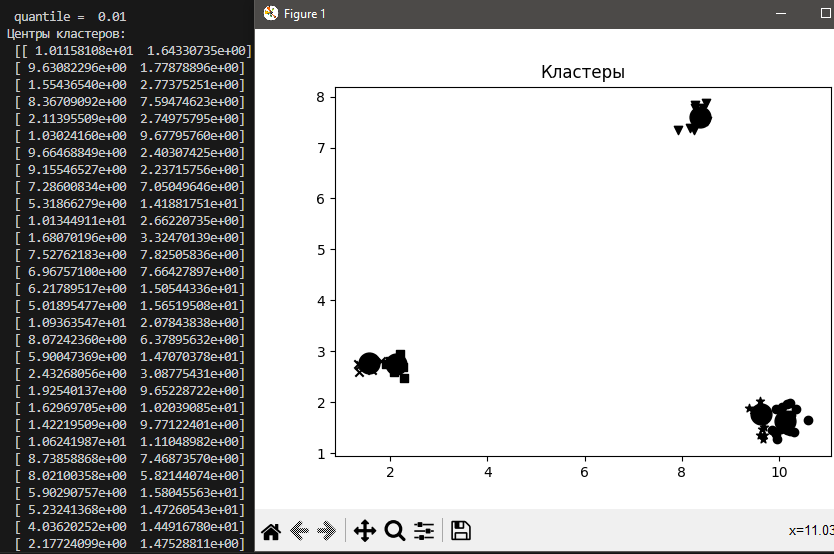
Задание 1

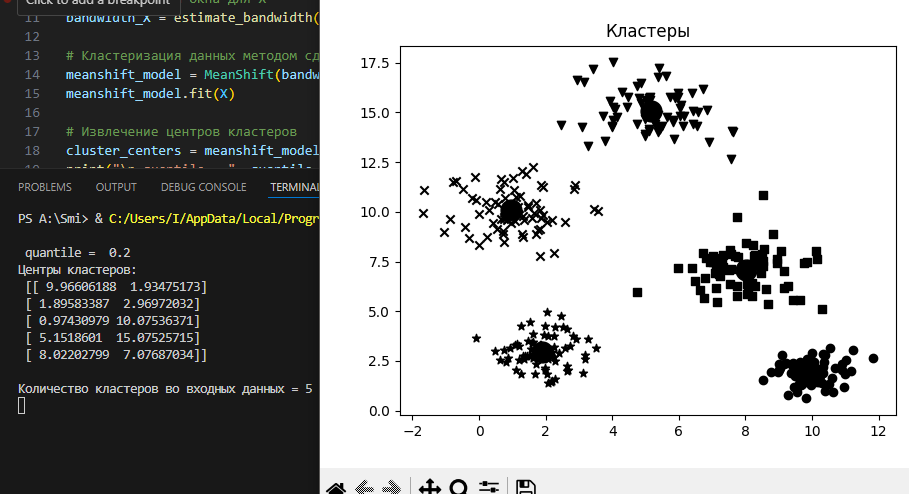
11)  


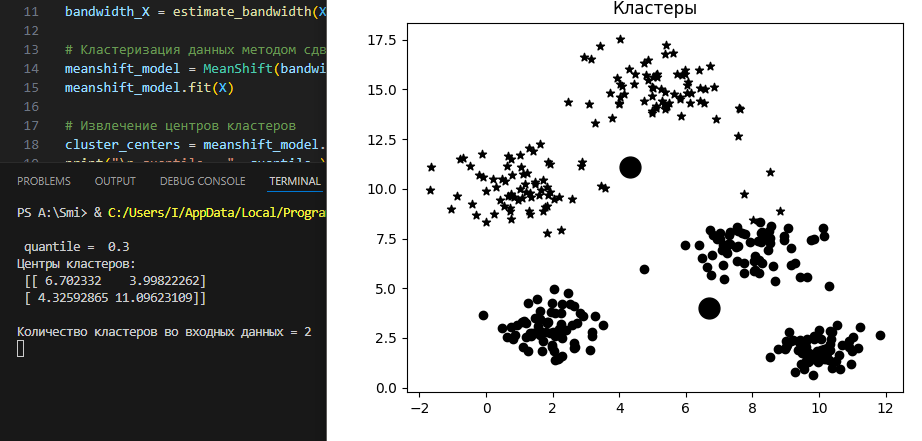
Задание 2

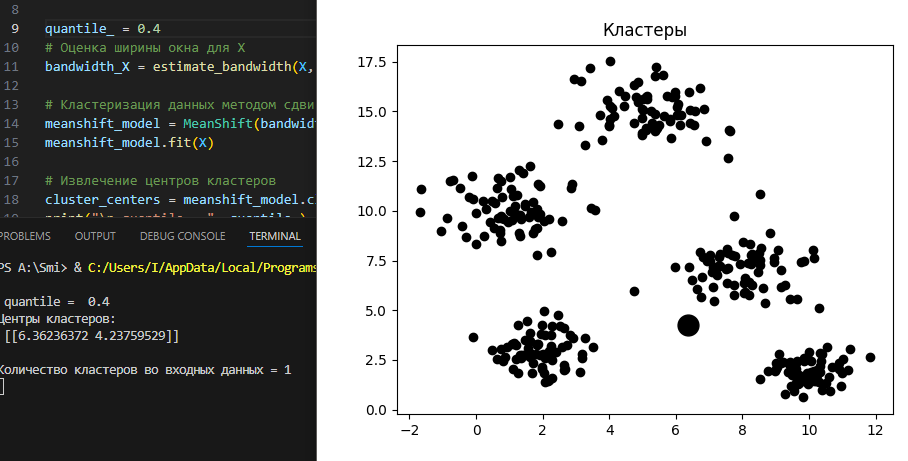
8)

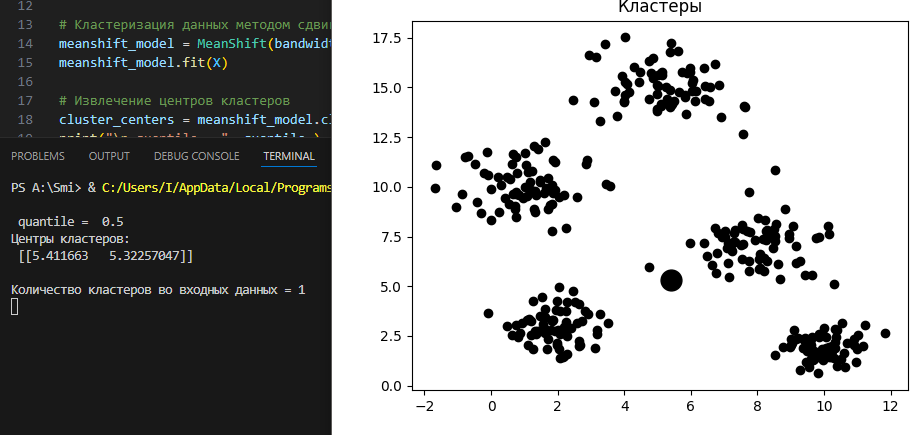


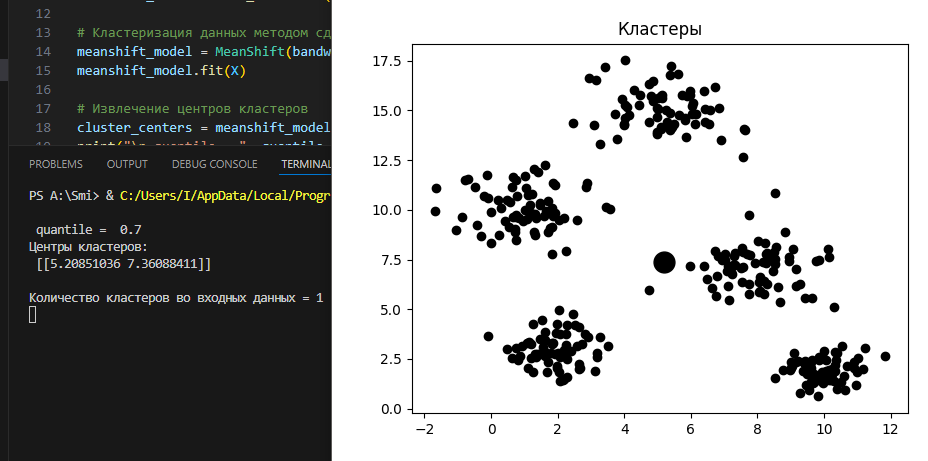


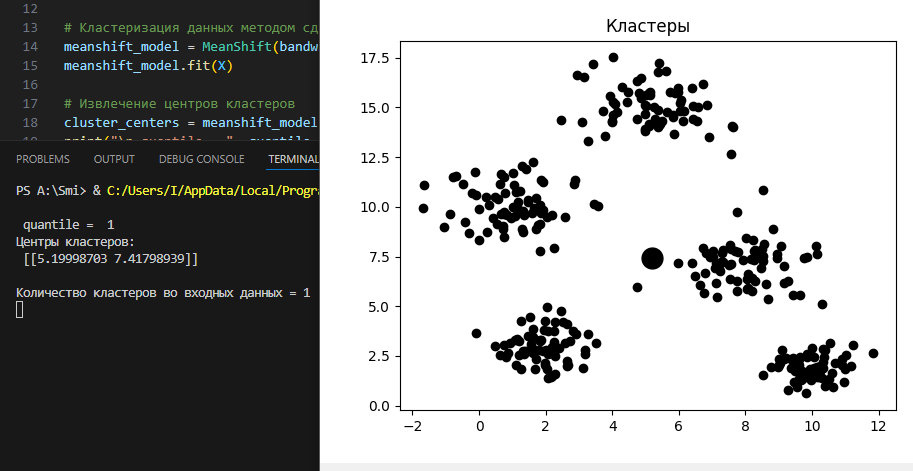


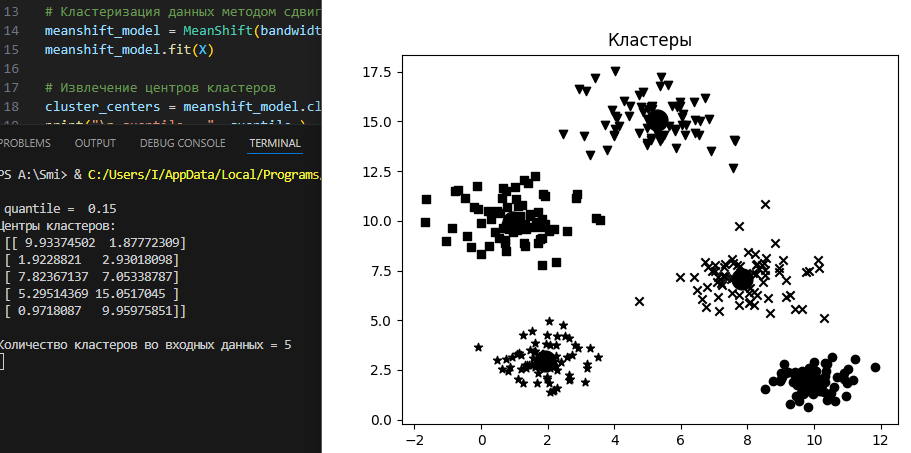


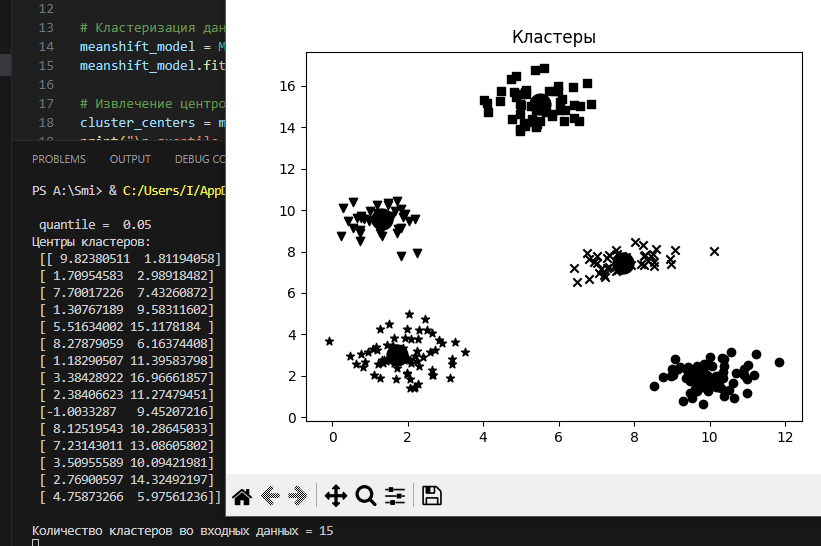












Низкие значения quantile (например, 0.01 или 0.05) приводят к переобучению и слишком большому количеству кластеров.

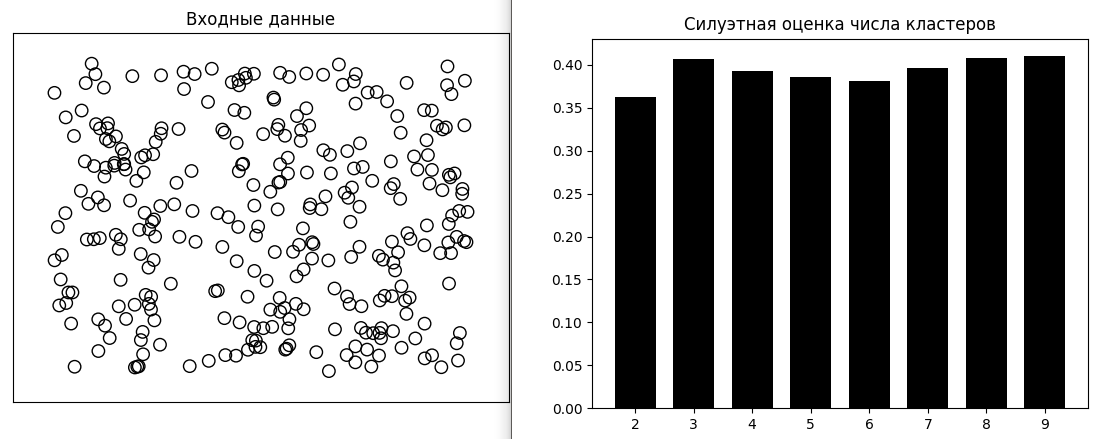
Средние значения quantile (например, 0.1 или 0.3) предлагают лучший баланс между количеством кластеров и их качеством, позволяя выявить более крупные и значимые группы.

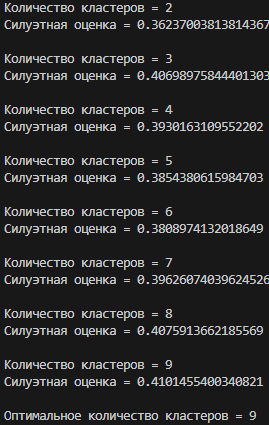
Высокие значения quantile (например, 0.5 и выше) сводят количество кластеров к минимуму, что может быть полезно для выделения единой группы, но может привести к потерям информации о структуре данных.

Таким образом, эффективность кластеризации зависит от выбора значения quantile. Для получения адекватных и интерпретируемых результатов рекомендуется использовать средние значения этого параметра, чтобы избежать как избыточного деления данных, так и слишком сильного их объединения.

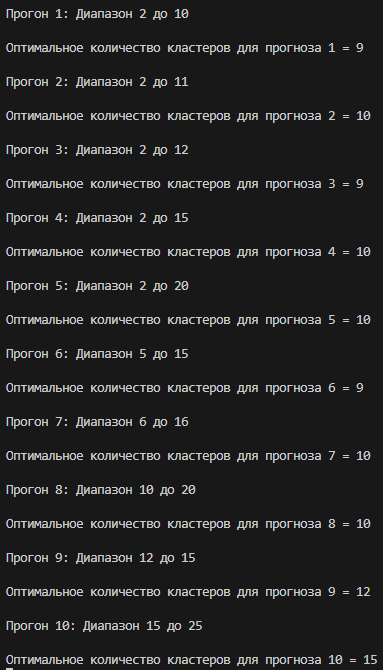
Задание 3

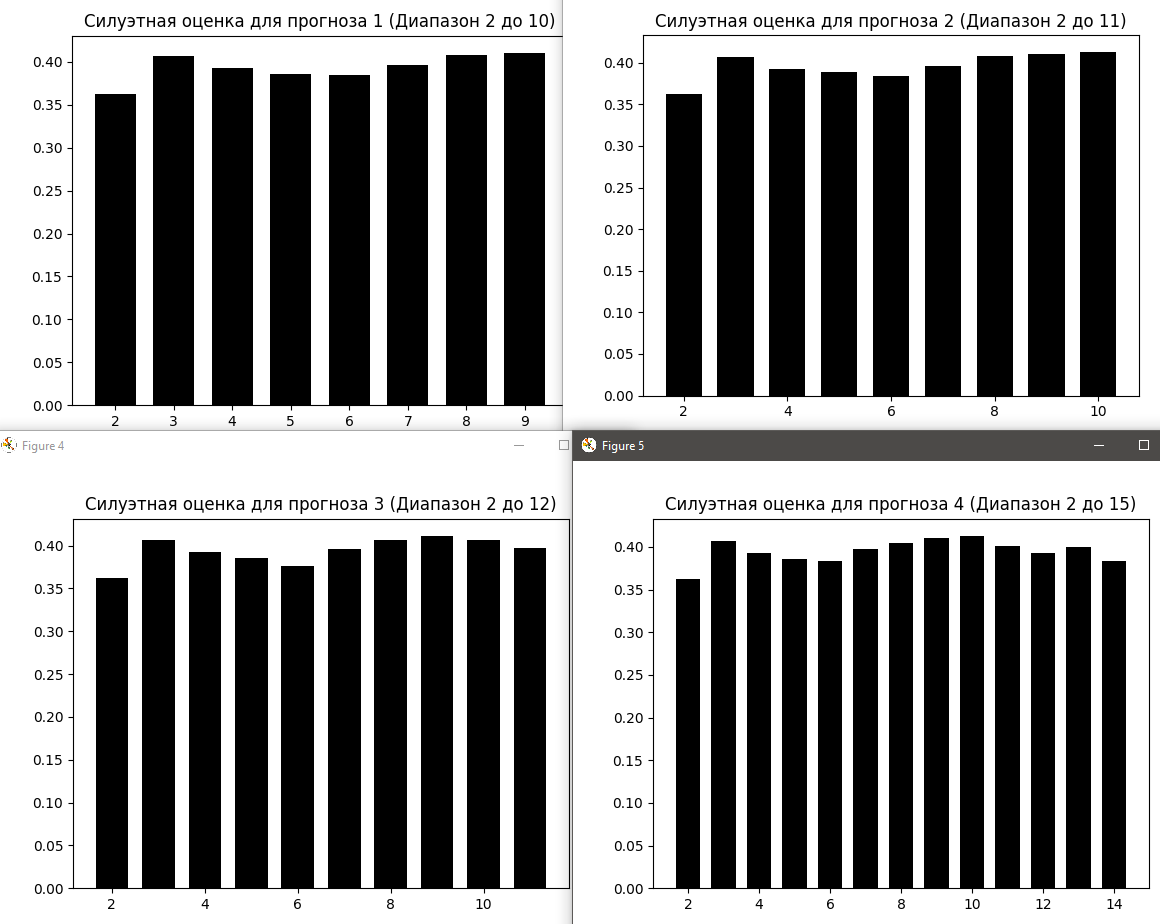
8)

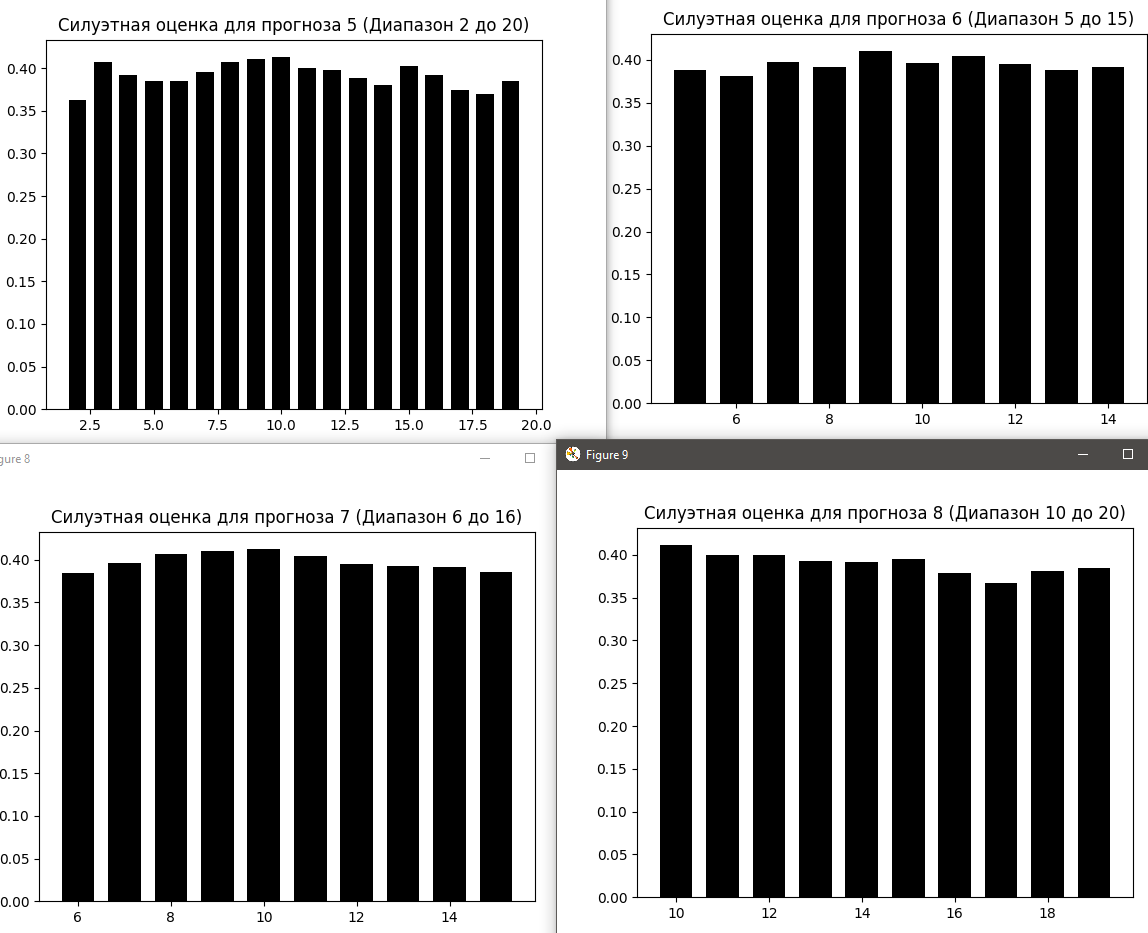


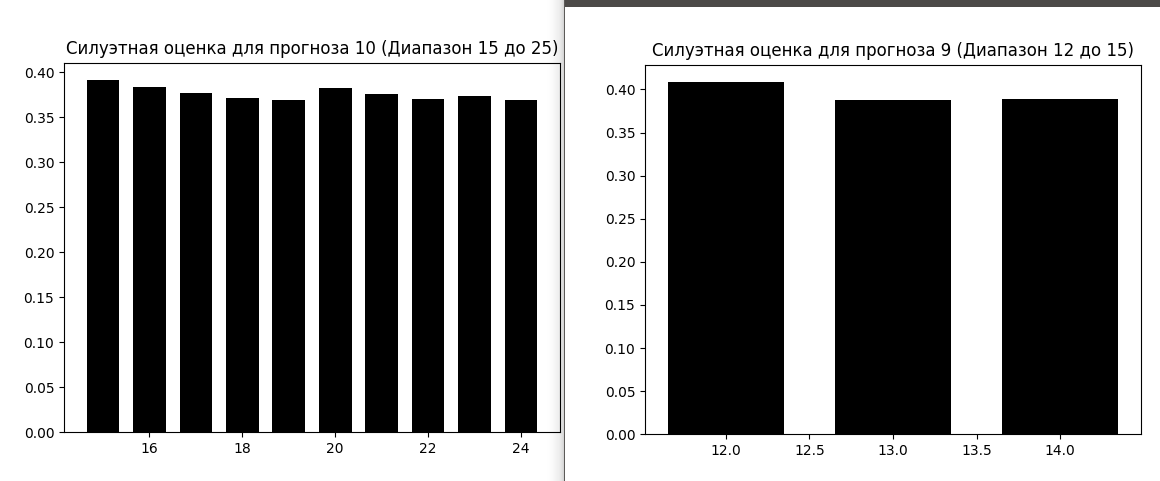


9)



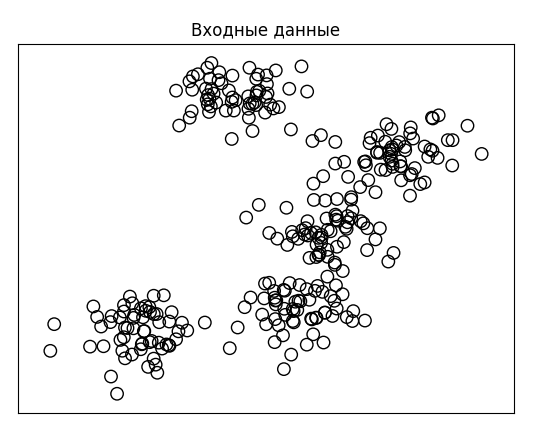


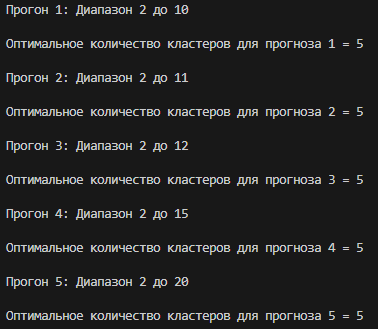


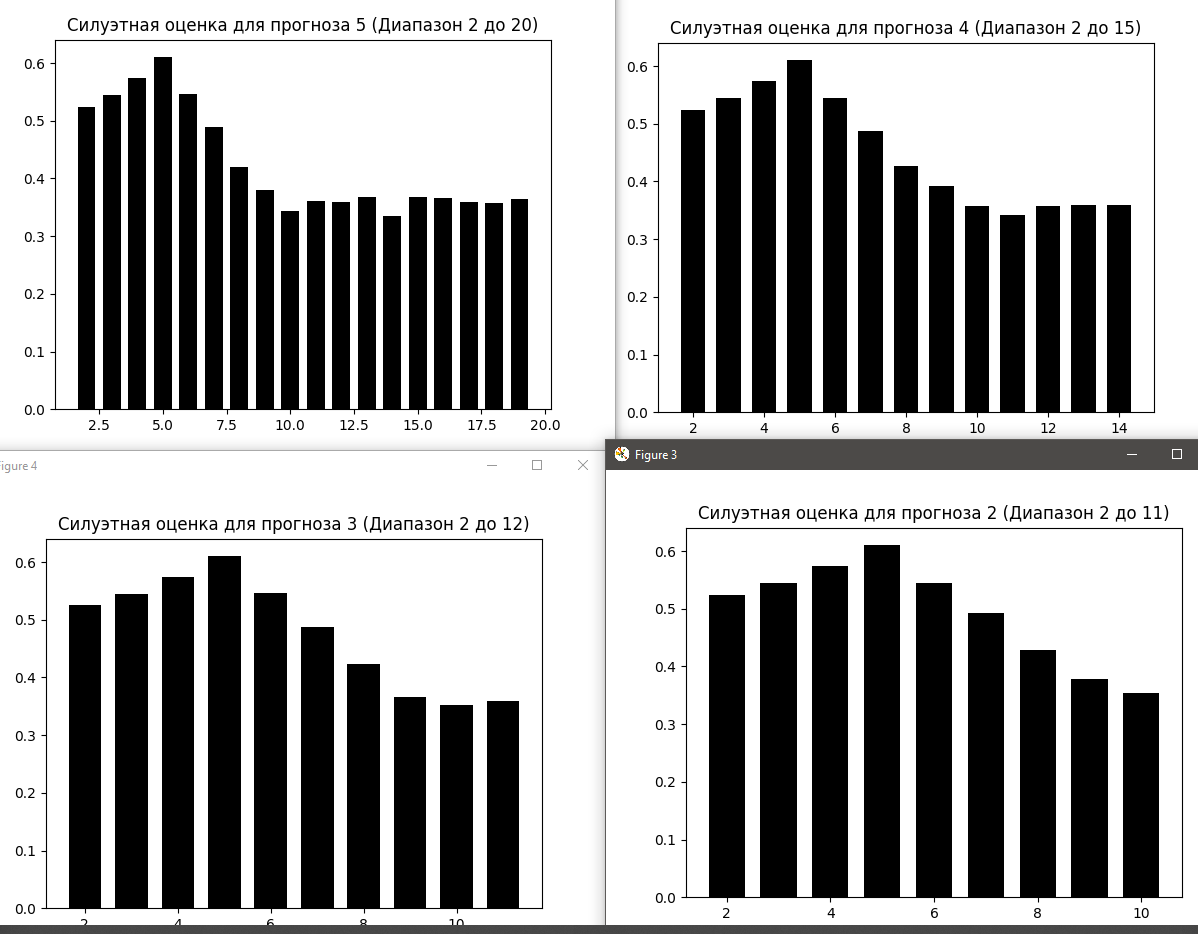


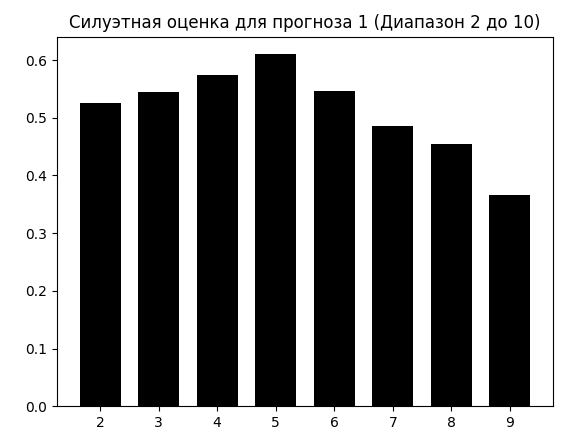
Результаты показывают, что оптимальное количество кластеров стабильно остаётся **10** в большинстве прогонов, кроме 9-го и 10-го, где оно значительно увеличивается до **12** и **19** соответственно. Это указывает на наличие более сложной структуры данных, требующей большего числа кластеров при расширении диапазона значений.

10)









Все 5 прогонов дали одинаковый результат — оптимальное количество кластеров равно 5. Увеличение верхней границы диапазона не повлияло на этот результат. Это подтверждает, что структура данных состоит из 5 кластеров, и алгоритм KMeans стабильно их выявляет, независимо от изменения диапазона.